CONFERENCING ADDITIONAL CALLERS INTO
AN ESTABLISHED VOICE BROWSING SESSION

Inventor(s):

Joseph Celi, Jr.

Peeyush Jaiswal

Victor S. Moore

International Business Machines Corporation

IBM Docket No. BOC9-2001-0043

IBM Disclosure No. BOC8-2001-0068

Express Mail Label No. EL 649719828 US

WP066313;1

BACKGROUND OF THE INVENTION

Technical Field

The present invention relates to multiplexed telephonic communications, and more particularly, to multi-party voice browsing sessions.

Description of the Related Art

The Voice Extensible Markup Language (VoiceXML) is a markup language useful in the creation and management of computer-human dialogs. VoiceXML capitalizes on existing telephony, markup language and data communications network technologies to facilitate user interaction with various speech-enabled services. In consequence, VoiceXML brings the advantages of Web-based development and content delivery to interactive voice response applications.

For example, VoiceXML can support features such as the playback of synthesized speech, digitized speech or audio, speech recognition, recognition of dual tone multi-frequency (DTMF) key input, as well as the recording of a speech input. Moreover, since VoiceXML leverages existing telephony and Internet technologies, no specialized access device is required for accessing online content. Thus, a caller can initiate a voice browsing session with a VoiceXML application and can interact with the VoiceXML application to perform selected tasks or to access online content.

Although VoiceXML provides voice application developers with access to a substantial feature set, several deficiencies have been associated with the conventional deployment of VoiceXML solutions. For instance, presently, the VoiceXML specification does not provide a mechanism in which additional callers can access an established

voice browsing session. Rather, to conference additional callers into an established

voice browsing session, specialized hardware and software can be required.

SUMMARY OF THE INVENTION

The invention disclosed herein provides a method for conferencing additional callers into an established voice browsing session. In particular, additional callers can be added to an existing voice browsing session using technology disposed at the application level rather than at the hardware level. When additional callers have been added to the established voice browsing session, requests originating from all conference participants, can be processed by the voice browser handling the voice browsing session. Moreover, the conference participants both can communicate with one another over a voice link, and also can receive audio from the voice browser.

One aspect of the present invention can include a voice browser based call conferencing method. The method can include establishing a voice browsing session between a calling party and the voice browser. An inbound call can be received from an additional party and the additional party can be conferenced into the voice browsing session. The conferencing step also can include conferencing selected ones of a plurality of additional parties into the voice browsing session. In any case, the conference can provide a voice communications link between the calling party and the additional party.

The conferencing step can include determining whether the inbound call is associated with an active voice browsing session and, if so, routing the inbound call to the voice browser associated with the active voice browsing session. Also, a determination can be made as to whether the voice browser associated with inbound call is configured for multiple callers. The conferencing step further can include

aggregating a voice data stream of the additional party with a voice data stream of the calling party into a single voice data stream and sending the single voice data stream for processing to the voice browser. Also, audio from the voice browser can be sent to the calling party and the additional party.

BRIEF DESCRIPTION OF THE DRAWINGS

There are shown in the drawings embodiments which are presently preferred, it being understood, however, that the invention is not limited to the precise arrangements and instrumentalities shown, wherein:

FIG. 1 is a block diagram of a speech processing system which has been configured to handle a voice browsing session in accordance with the inventive arrangements disclosed herein; and,

FIG. 2 is a flow chart illustrating the call conferencing method of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

The invention disclosed herein provides a method for conferencing additional callers into an voice browsing session which has been established between an initial caller and a voice browser. Specifically, the dialed number identification service (DNIS) of selected inbound calls which have been directed to a particular voice server can be used to route the inbound calls to an available and associated voice browser or voice browser application. Accordingly, as additional calls are received, the DNIS within each call can be identified and compared with active voice browsers to determine whether one of the active voice browsers has been configured to accept inbound calls for that DNIS. If so, the inbound call can be conferenced into an established voice browsing session.

Upon successfully adding the inbound call to the established voice browsing session, the voice data stream of the additional caller can be coordinated with the voice data stream of the initial caller. For example, audio from the various conference call participants can be aggregated into a single voice data stream that can be directed to an appropriate voice browser or voice browser application. Additionally, audio from the voice browser can be provided to each conference call participant or to selected participants. In any case, subsequent to adding the additional callers to the existing voice browsing session, requests originating from any of the conference call participants can be processed by the voice browser. In this manner, unlike prior art attempts at voice conferencing, additional callers can be conferenced into an

established voice browsing session at the application level without requiring additional

hardware.

FIG. 1 is a block diagram of a speech processing system which has been

configured to handle voice conferencing in accordance with the inventive arrangements.

5   Referring to FIG. 1, the speech processing system can include a public switched

telephone network (PSTN) 205, a computer communications network 208, a gateway

210, a voice server 215, and an application server 235. Callers can initiate telephone

calls or other voice communications via telephones 202 and 203 which are

communicatively linked to the PSTN 205. Although telephones 202 and 203 are shown

10   in FIG. 1, any of a variety of communications devices capable of functioning as a

speech interface such as voice-enabled personal digital assistants, wireless

telephones, cellular telephones, or the like can be communicatively linked to the PSTN

205.

    The gateway 210 can be a voice over IP (VoIP) gateway, a voice extensible

15   markup language (VoiceXML) gateway/server, or other suitable networking device

which can communicatively link the computer communications network 208, for

example the Internet, with the PSTN 205. The voice server 215, which can be

communicatively linked to the computer communications network 208, can include a

telephony and media (TM) component 225, a voice data stream manager 220 and a

20   voice browser 230. The voice server 215 further can include multiple instances of the

TM component 225, the voice data stream manager 220, and the voice browser 230.

The Voice server 215 can be configured to provide automatic speech recognition, text-

to-speech, audio, as well as dual tone multi-frequency processing and generation.

Alternatively, one or more of these functions can be integrated within the gateway 210.

The TM component 225 can be configured to manage call control including

receipt of inbound calls as well as call setup, maintenance, and termination of calls

5   between one or more callers to be teleconferenced and included within a voice

browsing session of the voice browser 230. The TM component 225 can route audio

from the voice browser to one or more of the callers engaged in the conference call and

voice browsing session. For example, audio can be routed to selected individual callers

or to all callers.

10   The voice data stream manager 220 can coordinate the various voice data

streams between multiple callers engaged in the telephone conference call and

participating in the voice browsing session. The voice data stream manager 220 can be

configured to aggregate the various voice data streams from each caller engaged in the

conference call and voice browsing session, as well as route the aggregated data

15   stream to one or more selected application programs in the application server 235. The

voice data stream manager 220 can discriminate among the voice data streams within

the aggregated voice data stream to selectively route particular voice data streams to

selected application programs. Still, it should be appreciated that callers involved in the

conference call and voice browsing session can communicate among themselves via

20   one or more communications links established through the PSTN 205.

The voice browser 230 can process markup languages, such as VoiceXML,

which can specify computer-human dialogs. Accordingly, the voice browser 230 can

receive audio input and provide audio output. Notably, the voice browser 230 can be

configured to accept multiple simultaneous calls. A parameter can be included to

indicate whether the voice browser 230 has been configured in this manner. For

example, the voice browser 230 can be configured to process voice data streams from

5   multiple callers that have been aggregated into a single data stream by the data stream

manager 220.

The application server 235 can be communicatively linked to the voice server

215 as shown in FIG. 1. Alternatively, the application server 235 can be

communicatively linked through the computer communications network 208. The

10   application server 235 can include one or more applications, such as VoiceXML

applications 240, which can be executed by the voice browser 230. Those skilled in the

art will recognize that the application programs 240 can be included within the voice

server 215.

FIG. 2 is a flow chart illustrating a method 200 of teleconferencing additional

15   callers into an existing telephone call and voice browsing session in accordance with

the inventive arrangements disclosed herein. The method 200 can begin in a state

wherein an existing call has been established between a caller and a voice server. The

voice data stream from the caller can be directed the voice browser associated with the

DNIS of the existing call. Still, the voice server can include a plurality of active voice

20   browsers and voice browsing sessions involving different callers. In step 205, an

inbound call can be received. In step 210, the DNIS of the inbound call can be

identified. After completion of step 210, the method can continue to step 215.

In step 215, a determination can be made as to whether the DNIS of the inbound call is associated with an active voice browsing session and corresponding voice browser. If so, the method can continue to step 220. If not, however, the method can continue to step 230 where the inbound call can be processed in a normal manner. For example, a voice browser associated with the DNIS of the inbound call can be activated and the voice data stream of the telephone call can be directed to that voice browser.

In step 220, a determination can be made as to whether the voice browser identified as being associated with the DNIS of the inbound call has been configured to accept multiple simultaneous calls. If so, the method can continue to step 225. If not, the method can continue to step 230 where, again, the inbound call can be processed in a normal manner. Proceeding with step 225, the caller can be added to identified and active voice browsing session. For example, the voice data streams of the caller and the inbound caller can be aggregated into a single voice data stream to be routed to the voice browser associated with the identified voice browsing session. Additionally, audio from the voice browser can be selectively routed to the caller, the inbound caller, or both.

Still, it should be appreciated that more than one additional caller can be conferenced into the telephone call and the voice browsing session. Accordingly, audio from the voice browser can be selectively routed to one of a plurality of conference call participants as appropriate. Once one or more inbound callers have been added to an existing call and voice browsing session, the speech from the caller and the inbound callers can be provided to the voice browser for processing. Further, the caller and

inbound callers can communicate with one another through a voice link over the PSTN for example.

The present invention can be realized in software in a centralized fashion in one computer system, or in a distributed fashion where different elements are spread across several interconnected computer systems. Any kind of computer system or other apparatus adapted for carrying out the methods described herein is suited. A typical implementation can include a general purpose computer system with a computer program that, when being loaded and executed, controls the computer system such that it carries out the methods described herein.

The present invention also can be embedded in a computer program product, which comprises all the features enabling the implementation of the methods described herein, and which when loaded in a computer system is able to carry out these methods. Computer program in the present context means any expression, in any language, code or notation, of a set of instructions intended to cause a system having an information processing capability to perform a particular function either directly or after either or both of the following: a) conversion to another language, code or notation; b) reproduction in a different material form.

This invention can be embodied in other forms without departing from the spirit or essential attributes thereof. Accordingly, reference should be made to the following claims, rather than to the foregoing specification, as indicating the scope of the invention.